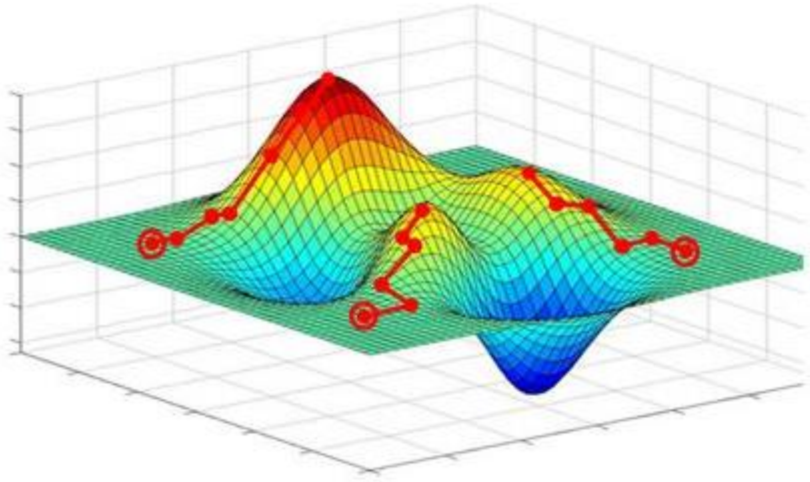


Introduction to Optimization in Computer Vision

EE, KAIST
김창익

Computational Imaging Lab.

Contents

- Introduction
  - Introduction to optimization
  - History
- Classification of optimization techniques in Computer Vision
  - Classification of optimization techniques
  - “Discrete” vs “Continuous”
  - Classification based on the nature of solution
  - Calculus of variations
  - Variational optimization
  - Other classification

Introduction

Introduction to Optimization

- Optimization refers to the class of problems that consists in choosing the best among a set of alternatives.
- *Best*, conveys a choice of criterion used to choose the solution and is usually expressed by means of a function that should be minimized or maximized.
- *Alternatives*, refers to the set of possible solutions or feasible set Ω.
- The "best" vector x^* is called the minimizer of $f(x)$ over Ω.

Optimization = Mathematical programming =
Energy minimization (in Image Understanding)

Introduction to Optimization

- Thus, the optimization problem is expressed as

$$\boldsymbol{x}^* = \min_{\boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x})$$
$$subject\,to\ \boldsymbol{x} \in \Omega.$$

- Vector $\boldsymbol{x} \in \boldsymbol{R}^n$ is an n-vector of independent variables.
- $f : \boldsymbol{R}^n \to \boldsymbol{R}$ is the objective function or cost function.
- The set Ω is a subset of $\boldsymbol{R}^n$ called the constraint set or feasible set.

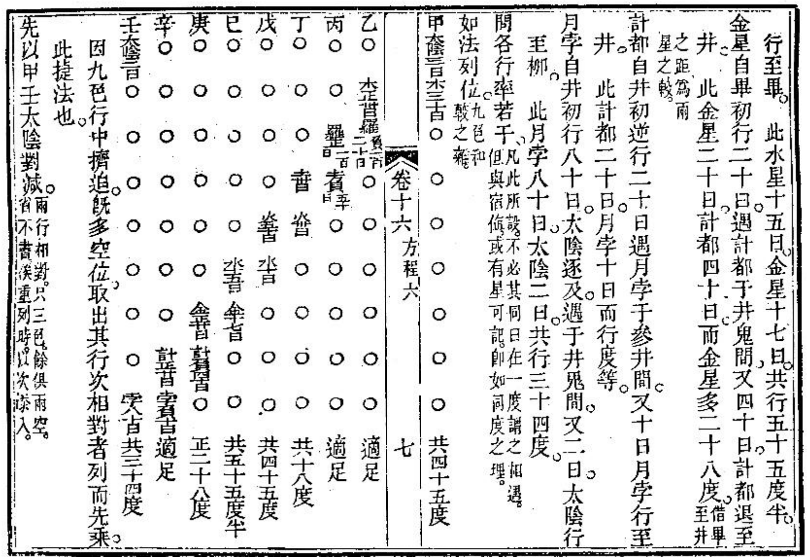
History

- Early stages: birth of linear algebra
  - Traced back to ancient China
  - Fangcheng (방정, rectangular arrays), used as early as BC300 to solve practical problems which amounted to linear systems
  - More specifically, in modern linear algebra there are two fundamentally different methods for solving systems of n linear equations in n unknowns: Gaussian elimination and determinants to solve

$$\boldsymbol{Ax} = \boldsymbol{y},\ \boldsymbol{A} \in \boldsymbol{R}^{nxn},\ \boldsymbol{A}\ invertible$$

History

- A fangcheng problem with 9 conditions in 9 unknowns



By Mei Wending, 1674

$$\left[\begin{array}{ccccccccc|c}
3 & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 45 \\
0 & 30 & -120 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 100 & -50 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 30 & 30 & 0 & 0 & 0 & 0 & 18 \\
0 & 0 & 0 & 0 & 50 & 10 & 0 & 0 & 0 & 45 \\
0 & 0 & 0 & 0 & 0 & 15 & 17 & 0 & 0 & \frac{111}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 20 & -40 & 0 & 28 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 & -10 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 80 & 34
\end{array}\right]$$

History

KAIST

- Optimization as a theoretical tool
  - Optimization problems crucial for theoretical mechanics and physics between the 17th and 19th centuries.
  - In the 1800s, Gauss built on early results in linear algebra for solving least-squares problems, which relied on solving an associated linear system.

$$\min_x \|Ax - y\|_2^2 \quad \Rightarrow \quad A^T(Ax - y) = 0. \quad \Rightarrow \quad x^* = (A^T A)^{-1} A^T y.$$

where $A \in \mathbf{R}^{m \times n}$, $y \in \mathbf{R}^m$ are given.

History

- Optimization as a theoretical tool
  - Lagrange was a key player, with the notion of duality which is central in optimization.

Given a nonlinear programming problem in standard form

$$\begin{aligned} \text{minimize} \ & f_0(x) \\ \text{subject to} \ & f_i(x) \leq 0, \ i \in \{1, \ldots, m\} \\ & h_i(x) = 0, \ i \in \{1, \ldots, p\} \end{aligned}$$

$\Rightarrow$

Lagrangian function Λ

$$\Lambda(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x).$$

History

- Advent of numerical linear algebra
  - With computers becoming available in the late 40s, numerical linear algebra ready to take off, motivated by the cold war effort.
  - Optimization played a key role in the development of linear algebra because it was an important source of applications and challenges
  - In the 70s, practical linear algebra linked to software, LINPACK and LAPACK written in FORTRAN, available in 80s
  - Followed by Matlab, Scilab, Octave, R, etc.

→ made linear algebra a commodity technology

History

- Advent of linear and quadratic programming
  - LP in the 40s for logistical problems arising in military operations
  - QPs popular in finance, where the linear term in the objective function refers to the expected negative return in an investment, and the squared term to risk introduced in 50s.
  - In the 60s-70s, a lot of attention devoted to nonlinear optimization problems, especially in the US in the 60s-80s thanks to large computers
  - In the Soviet union at that time, the focus was more towards optimization theory; due to restricted access to computing resources.

History

- Advent of convex programming
  - In the late 80s in Soviet Union, they discovered that the key property that makes an optimization problem "easy" is not linearity, but actually "convexity".
  - Roughly speaking, convex problems are easy, non-convex ones are hard.

History

- Present
  - A variety of fields, ranging from engineering design, statistics, and machine learning, to finance and structural mechanics.
  - Convex opt solvers: CVX (cvxr.com/cvx), YALMIP (users.isy.liu.se/johanl/yalmip/)
  - Large scale convex problems arising in machine learning, image processing, and so on.

Classification of Optimization Techniques (1)

Contents

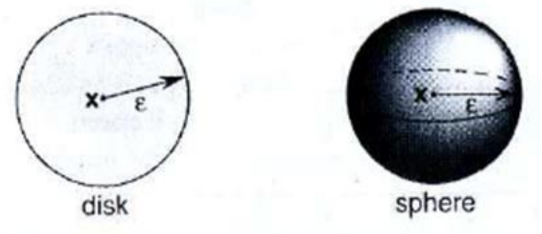
- Introduction
  - Introduction to optimization
  - History
- Classification of optimization techniques in Computer Vision
  - Classification of optimization techniques
  - “Discrete” vs “Continuous”
  - Classification based on the nature of solution
  - Calculus of variations
  - Variational optimization
  - Other classification

Classification of Optimization Techniques

- Optimization problems can be classified according to several criteria related to the properties of the objective function and also of the solution set S.
  - The nature of the possible solution set S
  - Constraints
  - The properties of the objective function f
- The most important classification is the one based on the nature of the solution set S, which leads us to classify optimization problems into following classes: continuous, discrete, combinatorial, and variational.

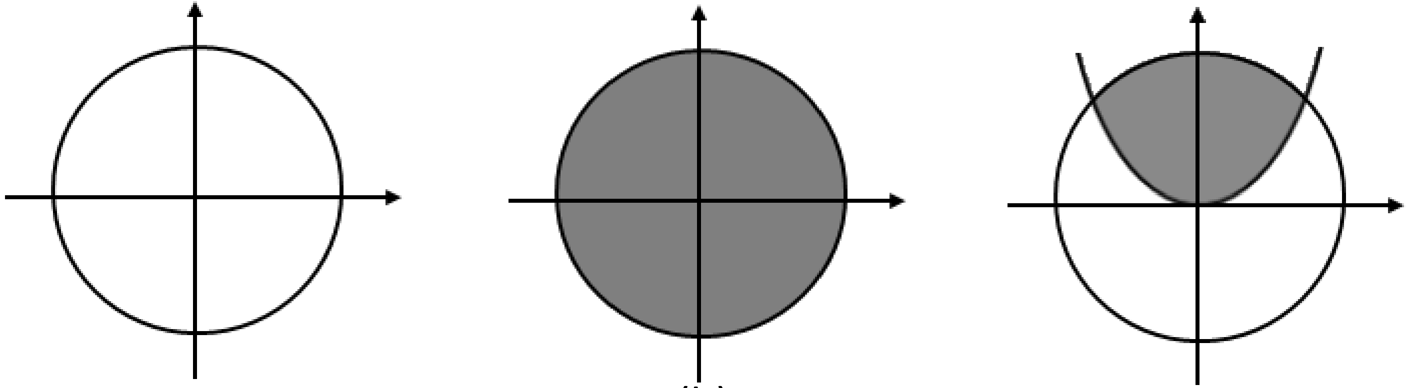
"Discrete" Vs. "Continuous"

- Definitions
  - A point x of a topological space S is an accumulated point, if and only if, for any open ball B_x, with $x \in B_x$, there exists an element $y \in S$ such that $y \in B_x$.
  - A point $x \in S$ which is not an accumulation point is called an isolated point of S. Thus x is isolated if there exists an open ball B_x, such that $x \in B_x$ and no other point of S belongs to B_x .
  - A topological space is discrete if it contains no accumulation points. It is continuous when all of its points are accumulation points.



Classification based on the nature of solution

- Continuous optimization problems
  - An optimization problem is called *continuous* when the feasible set S is a continuous subset of R^n.
  - The most common cases occur when S is a region of R^n .



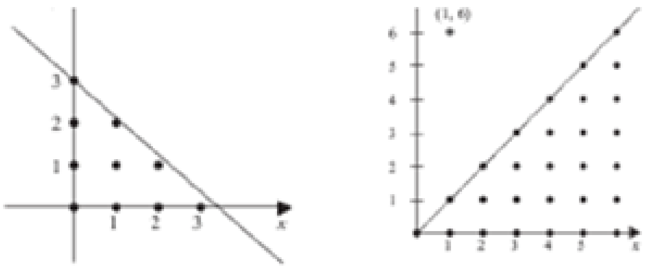
(a) (b) (c)

Classification based on the nature of solution

- Discrete optimization problems
  - An optimization problem is called *discrete* when the feasible set S is a discrete set. (i.e, S has no accumulation points)

$$S \subseteq Z^n = \{(i_1, i_2, \ldots, i_n); i_n \in Z\}.$$

  - The most common cases occur for



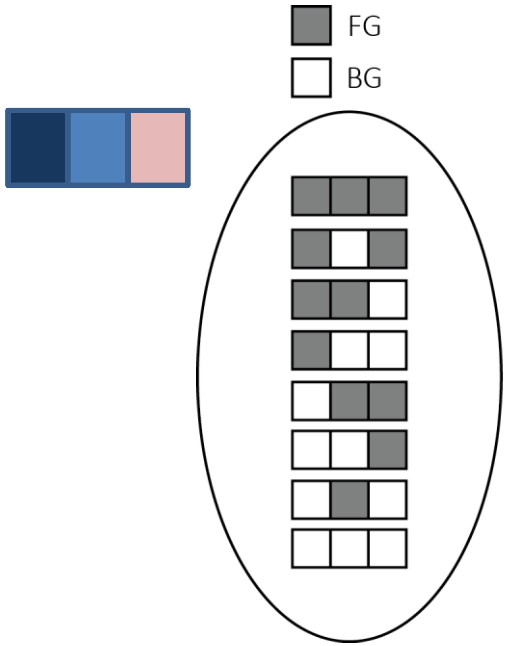
(a) (b)

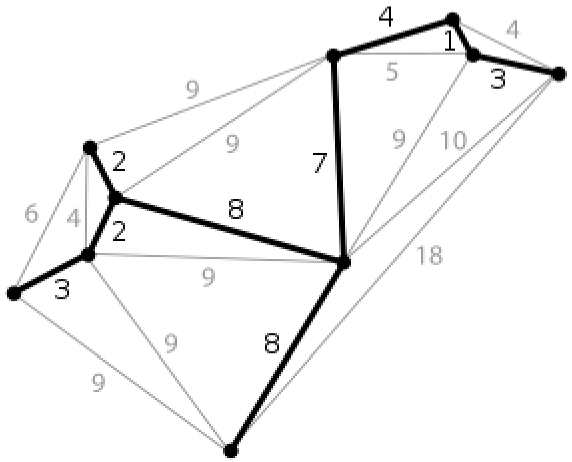
Classification based on the nature of solution

- Discrete optimization problems
  - As opposed to continuous optimization, the variables used in the mathematical program (or some of them) are restricted to assume only discrete values, such as the integers.
  - Two notable branches of discrete optimization are:
    1. combinatorial optimization, which refers to problems on graphs, matroids, and other discrete structures
    2. integer programming: LP relaxation, or cutting plane methods as exact algorithms
    - Note that these branches are intertwined.

Classification based on the nature of solution

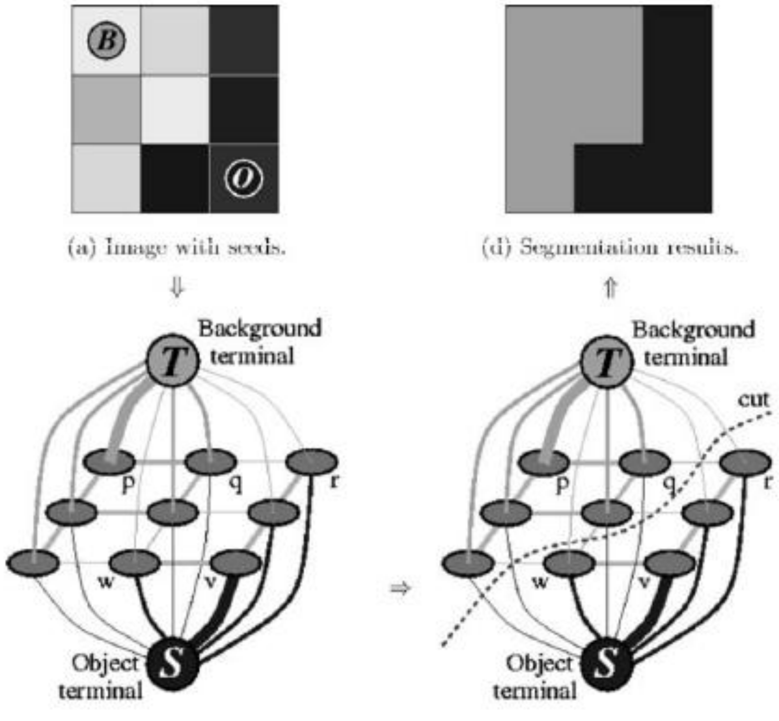
- Combinatorial optimization problems
  - In combinatorial optimization, the set of possible solutions S has a finite number of elements.
  - Any combinatorial optimization problem is also a discrete problem.
  - A solution can be expressed as a combination of representation of data. (i.e., it is finding an optimal object from a finite set of objects)





Classification based on the nature of solution

Object Segmentation (pixel labeling) via Graph Cuts



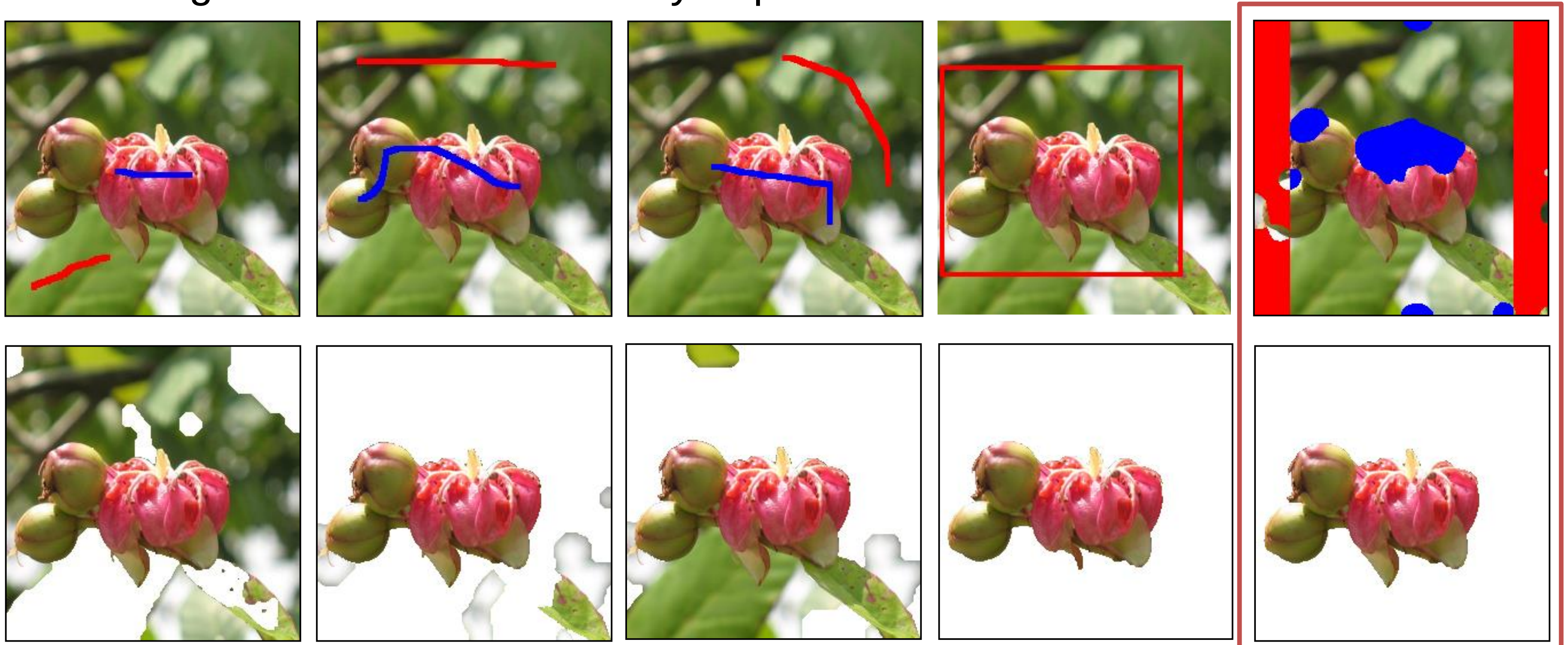
(b) Graph.

(c) Cut.

Classification based on the nature of solution

Ex1) Graph cut based segmentation

- ❖ Segmentation results heavily depend on user interaction.



Interactive Graph Cuts [1] | Interactive Graph Cuts [1] | Interactive Graph Cuts [1] | GrabCut [2] | Markov Chain Embedded Graph Cuts (proposed)

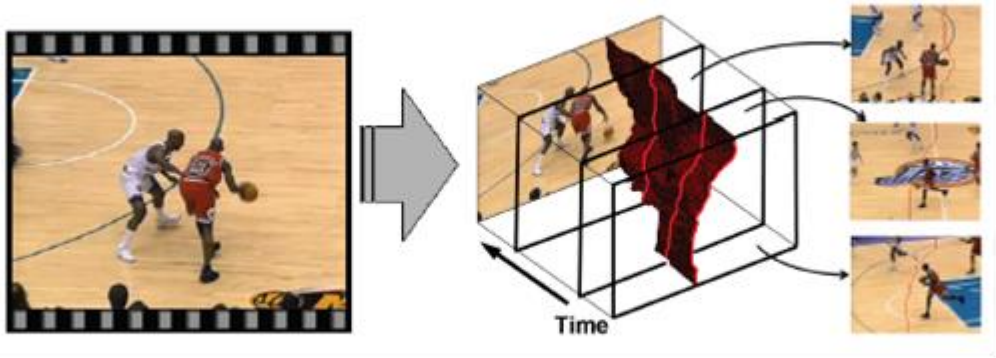
[1] Y. Boykov and G. Funka-Lea, “Graph cuts and efficient N-D image segmentation,” *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109-131, 2006.

[2] C. Rother, V. Kolmogorov, and A. Blake, “ “GrabCut” - Interactive Foreground Extraction using Iterated Graph Cuts,” *ACM Transactions on Graphics*, pp. 309-314, 2004.

[**proposed**] C. Jung, B.Kim, and C.Kim, "Automatic segmentation of salient objects using iterative reversible graph cut,” in *Proc. of IEEE International Conference on Multimedia & Expo (ICME)*, pp. 590-595, Singapore, July 2010.

Classification based on the nature of solution

Ex2) Video Retargeting using graph cuts



M. Rubinstein, A. Shamir, S. Avidan, “Improved seam carving for video retargeting,” ACM Transactions on Graphics, Vol. 27, No. 3, Aug, 2008.

Classification of Optimization Techniques (2)

Contents

- Introduction
  - Introduction to optimization
  - History
- Classification of optimization techniques in Computer Vision
  - Classification of optimization techniques
  - “Discrete” vs “Continuous”
  - Classification based on the nature of solution
  - Calculus of variations
  - Variational optimization
  - Other classification

Classification based on the nature of solution

- Variational Optimization Problems
  - An optimization problem is called a variational problem when its possible solution set (feasible set) S is an infinite dimensional subset of a space of functions (defined on a continuous domain).
  - In certain optimization problems, the unknown optimal solution might not be a number or a vector, but rather a continuous quantity, for example a function or the shape of a body.

Classification based on the nature of solution

- Variational Optimization Problems
  - Such a problem is also called an infinite-dimensional optimization problem, because a continuous quantity can not be determined by a finite number of certain degrees of freedom.
  - Calculus of Variation is used for solving such optimization problems.

Calculus of Variations

- A functional is a function of another function (or curve).
- A functional assigns a real number to each function (or curve) in some class.
- Functionals have an important role in many problems arising in analysis, optimization, mechanics, geometry, etc.

- Generally, we could write a functional:

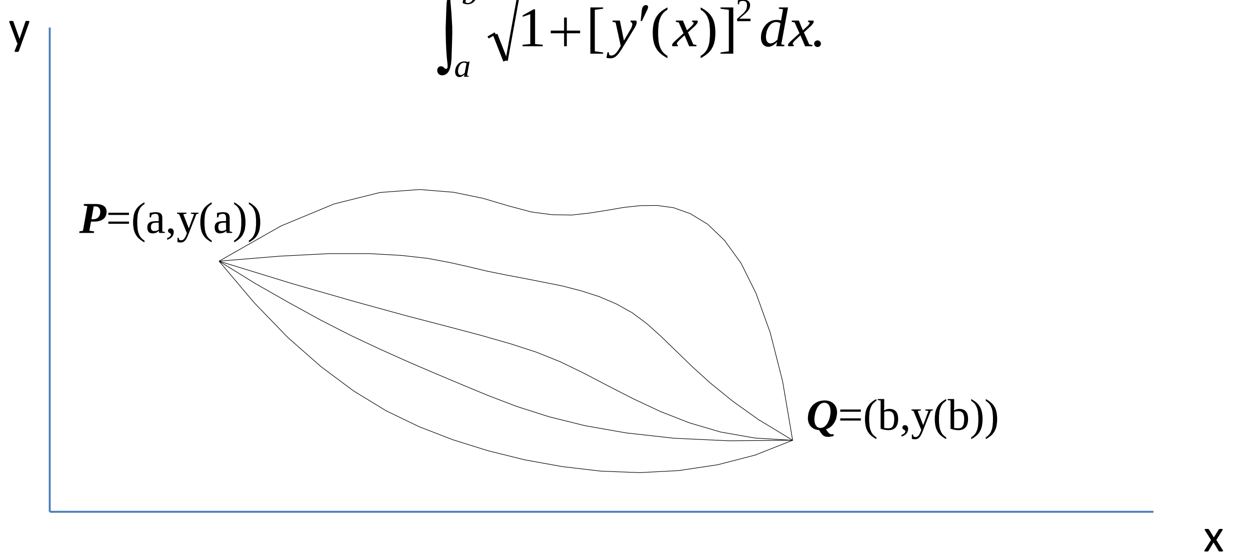
$$F[y(x)] = \int_a^b f(x, y(x), y'(x))dx.$$

- The function f in the integral is to be viewed as an ordinary function of the variables x, y, and y'.
- It is required that the function f in the integral have continuous partial derivatives of x, y, and y'.

Calculus of Variations

- A simple example. Let P and Q be two points in the xy-plane and consider the collection of all smooth curves which connect P to Q. Let y(x) be such a curve with P=(a,y(a)) and Q=(b,y(b)). The arc-length of the curve y(x) is given by the integral

$$\int_a^b \sqrt{1+[y'(x)]^2}\,dx.$$



Calculus of Variations

- Suppose the functional F obtains a minimum (or maximum) value.

  How do we determine the curve y(x)

  that produces such a minimum (or maximum) value for F?

Calculus of Variations

- We will see that the minimizing curve y(x) must satisfy a differential equation known as the *Euler-Lagrange Equation*.
- Theorem. If y(x) is a curve in $C_{[a,b]}$ which minimizes the functional

$$F[y(x)] = \int_a^b f(x, y(x), y'(x))dx.$$

then the following differential equation must be satisfied:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0.$$

This equation is called the Euler-Lagrange Equation.

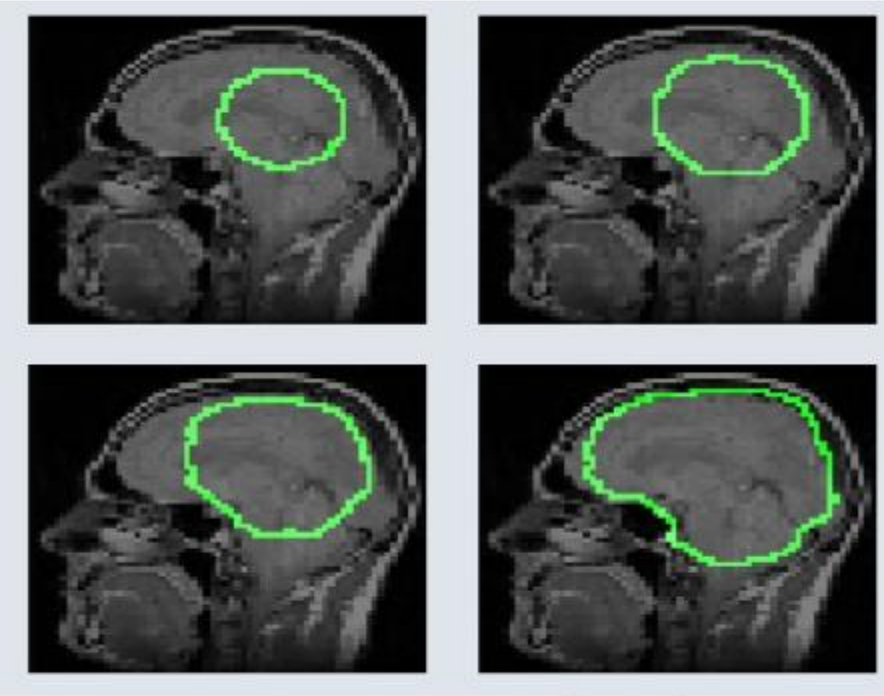
Calculus of Variations

Example *If* $F[y(x)] = \int_a^b \sqrt{1+[y'(x)]^2}\,dx$, *then the Euler-Lagrange Equation is given by:*

$$
\begin{aligned}
0 &= \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \\
&= 0 - \frac{d}{dx}\left(\frac{y'(x)}{\sqrt{1+[y'(x)]^2}}\right) \\
&= -\frac{\sqrt{1+[y'(x)]^2}\,y''(x) - [y'(x)]^2 y''(x)\left(1+[y'(x)]^2\right)^{-\frac{1}{2}}}{1+[y'(x)]^2} \\
&= -\frac{\left(1+[y'(x)]^2\right)y''(x) - [y'(x)]^2 y''(x)}{(1+[y'(x)])^{\frac{3}{2}}} \\
&= -\frac{y''(x)}{(1+[y'(x)])^{\frac{3}{2}}}
\end{aligned}
$$

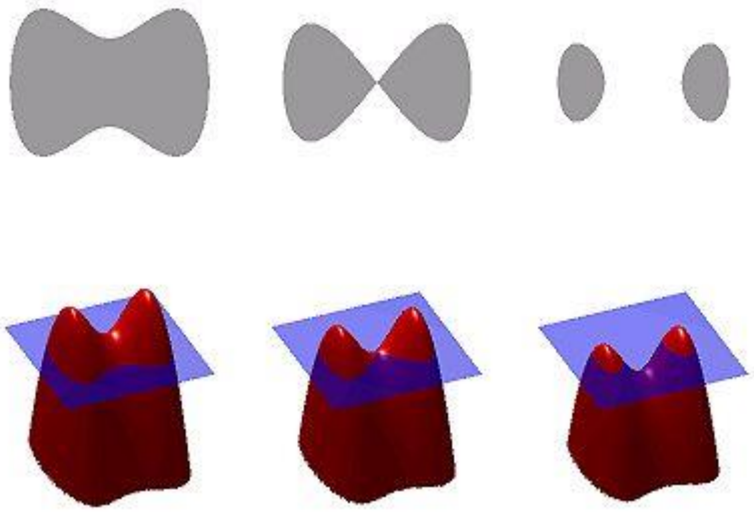
Variational optimization

- Finding curve functions



active contour model (snake)

$$E = \int_0^1 \frac{1}{2}\left[\alpha\left|x'(s)\right|^2 + \beta\left|x''(s)\right|^2\right] + E_{ext}(x(s))ds$$



Level set

Variational optimization

- Finding functions: Image denoising

$$\min_{u}\left\{\int_{\Omega}|\nabla u|dx+\frac{\lambda}{2}\int_{\Omega}(u-f)^2dx\right\}$$

Input Noisy Denoised

Other Classification

- Classification based on constraints
  - Unconstraint optimization
  - Constraint optimization
    - Equality constraints: $h_i(x) = 0, \quad i = 1, \dots, m.$
    - Inequality constraints: g(x)<0

Other Classification

- Classification based on the objective function
  - The properties of the objective function (or cost function) are also fundamental in order to devise strategies for the solution to optimization problems.

Other Classification

- Classification based on the objective function
  - Linear programming

    ex) Maximize -1000x + 5000y + 15000z,

    s.t. $x+2y+3z \leq 480$,

    $7x-3y-82 \geq 0$,

    $x, y, z \geq 0$.

  - Nonlinear programming

    Minimize $(x_1 - 3)^2 + (x_2 - 2)^2$

    Subject to ${x_1}^2 - x_2 - 3 \leq 0$

    $x_2 - 1 \leq 0$

    $-x_1 \leq 0$

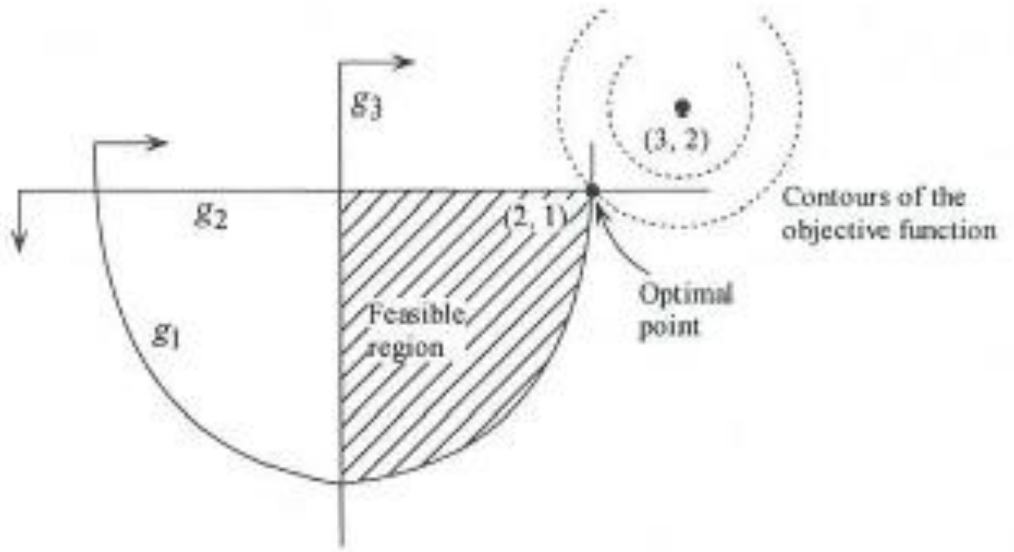


Figure 1.1 Geometric solution of a nonlinear problem.